

## Implementasi *Data Mining* Algoritma *Apriori* Dalam Menentukan Persediaan Susu Formula Bayi

**Desyanti**

Program Studi Informatika, Sekolah Tinggi Teknologi Dumai  
Jl. Utama Karya Bukit Batrem II  
Email: desyanti734@gmail.com

### ABSTRAK

Menambang data atau upaya untuk mengali suatu informasi dan pengetahuan data disebut dengan *Knowledge Discovery in Database (KDD)*. Salah satu algoritma yang paling populer dalam *data mining* adalah algoritma *apriori*. *Data mining* telah diimplementasikan keberbagai bidang diantaranya bisnis, perdagangan, pendidikan dan telekomunikasi. Di bidang bisnis misalnya hasil implementasi *data mining* menggunakan algoritma *apriori* dapat membantu para pebisnis dalam mengambil keputusan terhadap yang berhubungan dengan persediaan barang. Misalnya pentingnya sistem persediaan susu formula di sebuah mini market dan jenis susu formula apa yang menjadi prioritas utama yang harus distok untuk mengantisipasi kekosongan barang. Karena Minimnya stok barang dapat berpengaruh pada pelayanan konsumen dan pendapatan mini market, Oleh sebab itu ketersediaan jenis susu formula yang sering dibeli konsumen, mutlak untuk mendukung kelancaran pembelian susu formula sehingga aktivitas pelayanan konsumen berjalan dengan baik.

**Kata kunci:** *Data Mining*, Aturan Asosiasi, Algoritma *Apriori*.

### ABSTRACT

*Mapping data or attempts to multiply an information and data knowledge is called Knowledge Discovery in Database (KDD). One of the most popular algorithms in data mining is the a priori algorithm. Data mining has been implemented in various fields including business, trade, education and telecommunication. In the field of business such as the results of data mining implementation using a priori algorithm can help businessmen in making decisions against which related to inventory. For example the importance of the formula milk supply system in a mini market and what kind of formula milk is the main priority that must be stoked to anticipate the void of goods. Because of the lack of stock of goods can affect the consumer service and mini market income, Therefore the availability of the type of formula that is often purchased by consumers, absolutely to support the smooth purchase of formula milk so that customer service activity goes well.*

**Keywords:** *Data Mining*, Association Rules, *Apriori Algorithms*.

## Pendahuluan

Semua orang sependapat bahwa ASI (Air Susu Ibu) menjadi asupan terbaik dari ibu untuk bayi. Tetapi untuk beberapa kasus, banyak ibu yang tidak memberikan ASI karena alasan tertentu. Oleh karena itu demi memastikan sang bayi mendapatkan proses kembang secara sempurna, diperlukan asupan nutrisi tambahan atau alternatif. Susu Formula menjadi pilihan yang tepat bagi para ibu yang tidak bisa sepenuhnya memberikan asupan berupa ASI kepada sikecil. Susu formula banyak sekali dijual di apotik dan mini market terdekat, tidak terkecuali mini market sardini yang beralamat di Jl. Soekarno Hatta Bagan Besar Kota Dumai. Setiap hari banyak jenis produk yang terjual, diantaranya kebutuhan pokok sehari-hari termasuk susu formula untuk bayi. Karena banyaknya jenis susu formula yang dijual dan tidak adanya sistem komputerisasi persediaan barang pada toko tersebut mengakibatkan penumpukan stok susu formula, bahkan ada beberapa produk susu formula yang kurang laku terjual, sehingga menyebabkan perputaran modal penjualan menjadi tidak stabil. Melihat kondisi seperti itu diperlukan suatu analisis data untuk membantu pemilik toko dalam melakukan penyetokan susu formula agar tidak terjadi penumpukan stok susu .

*Data Mining* adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis (Fajar Astuti Hermawati, 2013). *Data mining* adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual (Kusrini & Luthfi, 2009). Hal terpenting dalam teknik *data mining* adalah aturan yang menentukan pola frekuensi yang tinggi antar himpunan item set yang disebut fungsi *association rule* (Aturan Asosiasi). Belakangan ini data mining telah diimplementasikan keberbagai bidang, diantaranya dalam bidang bisnis atau perdagangan , bidang pendidikan dan telekomunikasi. misalnya pentingnya sistem persediaan barang di suatu apotik dan jenis barang yang akan menjadi prioritas utama yang harus distok untuk mengantisipasi kekosongan barang. Karena minimnya stok barang dapat berpengaruh pada pelayanan konsumen dan pendapatan apotik.

Algoritma *apriori* termasuk jenis aturan asosiasi pada *data mining*. Selain apriori, yang termasuk pada golongan ini adalah metode *Generalized Rule Induction* dan *Algoritma Hash Based*. Aturan yang menyatakan asosiasi antara beberapa atribut sering disebut *affinity analysis* atau *market basket analysis*. Analisis asosiasi atau *association rule* adalah teknik *data mining* untuk menemukan aturan asosiatif antara suatu kombinasi *item* (Kusrini, 2009). Algoritma *apriori* menggunakan *knowledge* mengenai *frequent itemset* yang telah diketahui sebelumnya, untuk memproses informasi selanjutnya. Pada algoritma *apriori* untuk nmenentukan kandidat-kandidat yang mungkin muncul dengan cara memperhatikan minimum *support* dan minimum *confidence*.

*Support* adalah nilai penunjang, atau persentase kombinasi sebuah *item* dalam *database*. Rumus *support* sebagai berikut:

$$\text{Support (A)} = (\text{jumlah transaksi mengandung A} / \text{Total transaksi}) \times 100\%. \quad (1)$$

Sedangkan *confidence* adalah nilai kepastian yaitu kuatnya hubungan antar item dalam sebuah apriori. *Confidence* bisa dicari setelah pola frekuensi munculnya sebuah item ditemukan. Rumus untuk menghitung *confidence* adalah sebagai berikut: Contoh misal ditemukan aturan  $A \rightarrow B$  maka:

$$\text{Confidence } P(B|A) = (\text{Total transaksi mengandung A dan B} / \text{transaksi mengandung A}) \times 100\% \quad (2)$$

Adapun dua proses utama yang dilakukan dalam algoritma apriori yaitu

1. *Join* (penggabungan)  
Pada proses ini setiap item set dikombinasikan dengan item yang lainnya sampai tidak terbentuk kombinasi lagi.
2. *Prune* (pemangkasan)  
Pada proses ini hasil dari item yang telah dikombinasikan tadi lalu dipangkas dengan menggunakan *minimum support* yang telah ditentukan oleh *user*. Dua proses utama tersebut merupakan langkah yang akan dilakukan untuk mendapat *frequent itemset*. Walaupun algoritma *apriori* mudah untuk dipahami dan diimplementasikan dibandingkan algoritma yang lainnya yang memang diterapkan untuk proses *association rule*, akan tetapi algoritma *apriori* juga memiliki kekurangan yaitu: untuk melakukan pencarian *frequent itemset*, algoritma *apriori* harus melakukan *scanning database* berulang kali untuk setiap kombinasi *item*. Hal tersebut membutuhkan banyak waktu yang dibutuhkan untuk melakukan *scanning database*. Selain itu dibutuhkan *generate candidate* yang besar untuk mendapatkan kombinasi *item* dari *database*.

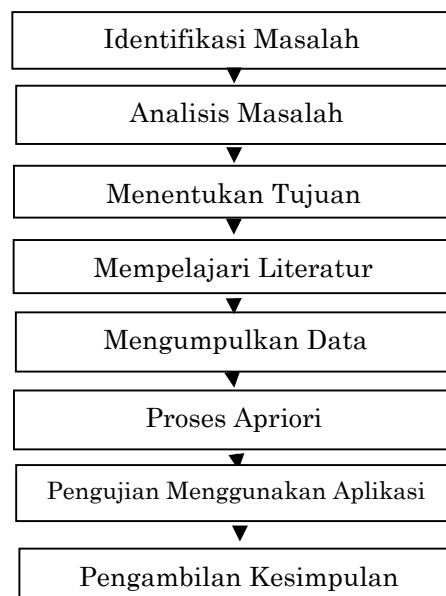
Pada penelitian sebelumnya Eka Novita Sari (2013) melakukan penelitian dengan judul Analisa Algoritma *Apriori* Untuk Menentukan Merek Pakaian Yang Paling Diminati Pada Mode *Fashion Group* Medan. Dalam penelitiannya membahas tentang analisa data menggunakan algoritma *Apriori* untuk mengetahui merek pakaian yang paling diminati oleh para konsumen pada *Mode Fashion Group* Medan. Data yang akan dianalisa diambil dari data transaksi yang berasal dari data penjualan perharinya. Setelah dilakukan perhitungan secara manual, kemudian dilakukan perhitungan secara komputerisasi dengan menggunakan bantuan *tool* Tanagra versi 1.4 dimulai dengan cara mengimport data penjualan perbulan dalam format tabular menggunakan *Microsoft Excel*. Semakin banyak data, maka semakin sulit dalam membuat format tabular karena harus menginput data satu persatu.

Sedangkan pada Jurnal Nugroho Wandu, Rully A. Hendrawan, dan Ahmad Mukhlason *Volume 1 – September 2012* ISSN 2301-9271 dengan judul “Pengembangan Sistem Rekomendasi Penelusuran Buku dengan Penggalian *Association Rule* Menggunakan Algoritma *Apriori* (Studi Kasus Badan Perpustakaan dan Kearsipan Provinsi Jawa Timur)”. Dalam penelitiannya membahas tentang penggunaan algoritma *Apriori* untuk menghasilkan rekomendasi buku dari transaksi peminjaman buku yang ada. Pada penelitian yang dilakukan, digunakan data transaksi dari tahun 2009 sampai 2012. Batas *minimum support* yang diberikan sebesar 1 hingga 4, dan

panjang *itemset* berbanding terbalik dengan jumlah minimum *support* yang diberikan, semakin besar nilai minimum *support* yang diberikan semakin kecil panjang jumlah *itemset* yang terbentuk.

### Metode Penelitian

Penelitian ini menggunakan Metode *apriori* dalam menentukan persediaan susu formula. Ada beberapa langkah-langkah yang harus dilalui agar penelitian ini bisa berjalan dengan baik. Kerangka kerja yang harus diikuti bisa dilihat pada Gambar 1.



**Gambar 1** Kerangka kerja

Berdasarkan Gambar 1 dijabarkan urutan-urutan langkah kerja sebagai berikut:

1. Identifikasi Masalah  
Masalah yang diidentifikasi dalam penelitian ini adalah
  - a. Tidak adanya sistem komputerasi persediaan barang pada mini market sarbini mengakibatkan penumpukan stok susu formula, bahkan ada beberapa produk susu formula yang kurang laku terjual, sehingga menyebabkan perputaran modal penjualan menjadi tidak stabil.
  - b. Belum adanya sebuah sistem yang menganalisis data susu formula untuk membantu pemilik toko dalam melakukan penyetokan susu formula agar tidak terjadi penumpukan stok susu.
2. Analisis Masalah  
Analisis masalah pada penelitian ini dilakukan dengan dua metode yaitu metode deskriptif dan metode komperatif.
  - a. Metode Deskriptif

Pada metode ini data yang ada dikumpulkan, disusun, dikelompokkan dan dianalisis sehingga diperoleh beberapa gambaran yang jelas pada masalah yang akan dibahas.

b. Metode Komperatif

Pada metode ini analisis dilakukan dengan cara membandingkan teori dan praktek sehingga diperoleh gambaran yang jelas tentang persamaan dan perbedaan di antara keduanya.

3. Menentukan Tujuan

Berdasarkan uraian di atas, tujuan yang akan dicapai dari penelitian ini adalah menghasilkan pola kombinasi *itemset* yang bisa dijadikan sebagai informasi yang sangat berharga dalam pengambilan keputusan untuk mempersiapkan stok susu formula.

4. Mempelajari Literatur

Literatur-literatur yang dipakai sebagai bahan referensi dalam penelitian ini adalah dari jurnal-jurnal ilmiah, modul pembelajaran dan buku tentang *data mining*. Literatur-literatur ini akan menjadi pedoman untuk melakukan penelitian agar memudahkan proses penelitian.

5. Mengumpulkan Data

Metode pengumpulan data dilakukan dengan cara mengambil data transaksi penjualan pada mini market sarbini selama tiga bulan terakhir.

6. Proses *Apriori*

Ada dua tahapan dalam melakukan perhitungan menggunakan algoritma *apriori*, yaitu yang pertama Analisa Pola Frekuensi Tinggi dan yang kedua Pembentukan Aturan Asosiasi.

a. Analisa Pola Frekuensi Tinggi

Mengumpulkan data awal dan dikelompokkan sesuai dengan transaksinya, kemudian menetapkan besarnya besaran  $\Phi$  dan nilai minimum *Support* dan *Confidence*.

b. Pembentukan Aturan Asosiasi

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum *confidence* yang telah ditentukan sebelumnya.

7. Pengujian Menggunakan Aplikasi Yang Ditentukan

Pada tahap ini, hasil dari analisis diuji kembali menggunakan sistem *data mining* yang sudah ada. *Tools* yang digunakan sebagai pengujian sistem adalah *XL Miner*.

8. Pengambilan Keputusan

Setelah diuji, hasil analisis antara *manual ways* (cara manual) dan pengujian memanfaatkan *tools* akan terlihat perbandingannya. Langkah berikutnya adalah penentuan atau pengambilan keputusan terhadap *knowledge* yang baru didapat yaitu berupa susu formula mana yang banyak diminati oleh konsumen sehingga dapat ditentukan persediaan stok susu formula.

## Hasil dan Pembahasan

Berdasarkan dari penelitian yang dilakukan penulis, berikut penerapan dari metode *apriori* yang digunakan yaitu pengetahuan antar jenis susu formula yang dijual. Jenis-jenis produk yang muncul bersamaan pada tiap transaksi dapat menjadi masukan penting dalam melakukan usaha peningkatan penjualan tersebut untuk mengetahui susu formula mana yang paling diminati konsumen. Algoritma *apriori* aturan asosiasi yang merupakan suatu tugasnya untuk menemukan atribut-atribut yang terjadi bersamaan. Tugas asosiasi mencoba menemukan aturan untuk menguantifikasi hubungan antara dua atau lebih atribut, bersama-sama dengan ukuran *support* dan *confidence* yang memenuhi syarat.

Ada dua tahapan dalam melakukan perhitungan menggunakan algoritma *apriori*, yaitu yang pertama Analisa Pola Frekuensi Tinggi dan yang kedua Pembentukan Aturan Asosiasi.

### 1. Analisa Pola Frekuensi Tinggi

Data yang digunakan adalah data jenis susu formula bayi dan data transaksi penjualan susu formula bayi pada mini market sarbini, Dalam melakukan perhitungan manual pada pembahasan ini, penulis mengambil 7 jenis data susu formula dan 14 sampel data dari transaksi yang dapat dilihat pada Tabel 1 dan Tabel 2.

**Tabel 1** Jenis produk susu formula

No	Jenis Susu Formula
1	SGM
2	Bebelac
3	BMT
4	Lactogen
5	Batita
6	Dancow
7	Enfamil

Data tersebut meliputi item-item susu formula bayi yang dibeli konsumen Data selengkapnya tampak pada Tabel 2.

**Tabel 2** Transaksi item belanja susu formula bayi

No	Jenis Susu Formula
1	BMT, Batita, Lactogen
2	SGM, Dancow, Lactogen
3	Laktogen, Enfamil, Dancow
4	Batita, Lactogen, Enfamil, Bebelac
5	Bebelac, SGM, BMT
6	Dancow, SGM, Bebelac, Enfamil
7	Enfamil, Lactogen
8	BMT, Enfamil, Batita
9	Dancow, SGM, Enfamil

10	Bebelac, Lactogen
11	Batita, BMT, Bebelac, Dancow
12	SGM, Bebelac, Dancow
13	Dancow, Lactogen, SGM, Bebelac
14	Lactogen, Batita, Enfamil, Bebelac, BMT

Kemudian menetapkan besaran ( $\Phi$ ), nilai minimum besaran *support* dan besaran *confidence* yang diinginkan untuk dipenuhi oleh aturan asosiasi yang ingin dihasilkan yaitu:

$$\Phi = 4, \text{Support} = 30\%, \text{Confidence} = 70\%$$

Kemudian pada tabel 3 disusun *itemset* yang memiliki frekuensi *itemset* minimum besaran  $\Phi=4$  yang telah ditetapkan. Akan dimulai dengan membahas setiap 1-*itemset*.

**Tabel 3** Jumlah kemunculan 1-*item set*

No	1- <i>itemset</i>	Support
1	SGM	6
2	Bebelac	10
3	BMT	5
4	Lactogen	8
5	Batita	5
6	Dancow	7
7	Enfamil	6

{SGM}, {Bebelac}, {BMT}, {Lactogen}, {Batita}, {Dancow} dan {Enfamil} *itemset* ini berhasil muncul melebihi 4 kali. Di lanjutkan dengan membahas setiap 2-*itemset* pada Tabel 4.

**Tabel 4** Jumlah kemunculan 2-*itemset*

No	2- <i>Itemset</i>	Support
1.	SGM, Bebelac	5
2.	SGM, BMT	1
3.	SGM, Lactogen	2
4.	SGM, Batita	0
5.	SGM, Dancow	5
6.	SGM, Enfamil	1
7.	Bebelac, BMT	3
8.	Bebelac, Lactogen	5
9.	Bebelac, Batita	3
10.	Bebelac, Dancow	6
11.	Bebelac, Enfamil	4

12.	BMT, Lactogen	2
13.	BMT, Batita	4
14.	BMT, Dancow	1
15.	BMT, Enfamil	2
16.	Lactogen, Batita	3
17.	Lactogen, Dancow	3
18.	Lactogen, Enfamil	4
19.	Batita, Dancow	1
20.	Batita, Enfamil	3
21.	Dancow, Enfamil	2

Tabel 4 dapat dijadikan sebagai penelaahan mengenai keseluruhan *2-itemset* tersebut akan membawa pada suatu kesimpulan bahwa hanya {SGM,Bebelac}, {SGM, dancow}, {Bebelac, laktogen}, {Bebelac,dancow}, {Bebelac, Enfamil},{BMT, batita} dan {laktogen, Enfamil} yang merupakan *2-itemset* yang memenuhi besaran  $\Phi = 4$ , Selanjutnya menghitung besaran *support* dan *confident*.

$$Support (A) = (\text{jumlah transaksi mengandung A} / \text{Total transaksi}) \times 100\% \quad (1)$$

$$Confidence (B|A) = (\text{Total transaksi mengandung A dan B} / \text{transaksi mengandung A}) \times 100\% \quad (2)$$

Adapun hasil dari perhitungan *support* dan *confident* dapat dilihat pada Tabel 5.

**Tabel 5** Hasil perhitungan *support* dan *confident*

Data Item	<i>Support</i>	<i>Confident</i>
SGM, Bebelac	5/14 =35,7 %	5/6 = 83,3 %
		5/10 = 50 %
SGM, Dancow	5/14 =35,7 %	5/6 = 83,3 %
		5/7 = 71,4 %
Bebelac, Lactogen	5/14 =35,7 %	5/10 = 50 %
		5/8 = 62,5 %
Bebelac, Dencow	6/14 =42,8%	6/10 = 60 %
		6/7 = 85,7 %
Bebelac, Enfamil	4/14 =28,6 %	4/10 = 40 %
		4/6 = 66,6 %
BMT, Batita	4/14 =28,6%	4/5 = 80%
		4/5 = 80 %
Lactogen, Enfamil	4/14 =28,6%	4/5 = 80%
		4/6 = 66,6 %

2. Pembentukan Aturan Asosiasi

Dari Tabel 5 dapat dilihat hasil perhitungan *support* dan *confident*, maka dilakukan seleksi dari aturan asosiasi yang memenuhi nilai minimum (*Support*) dan minimum (*Confidence*) saja. Adapun aturan asosiasi yang memenuhi

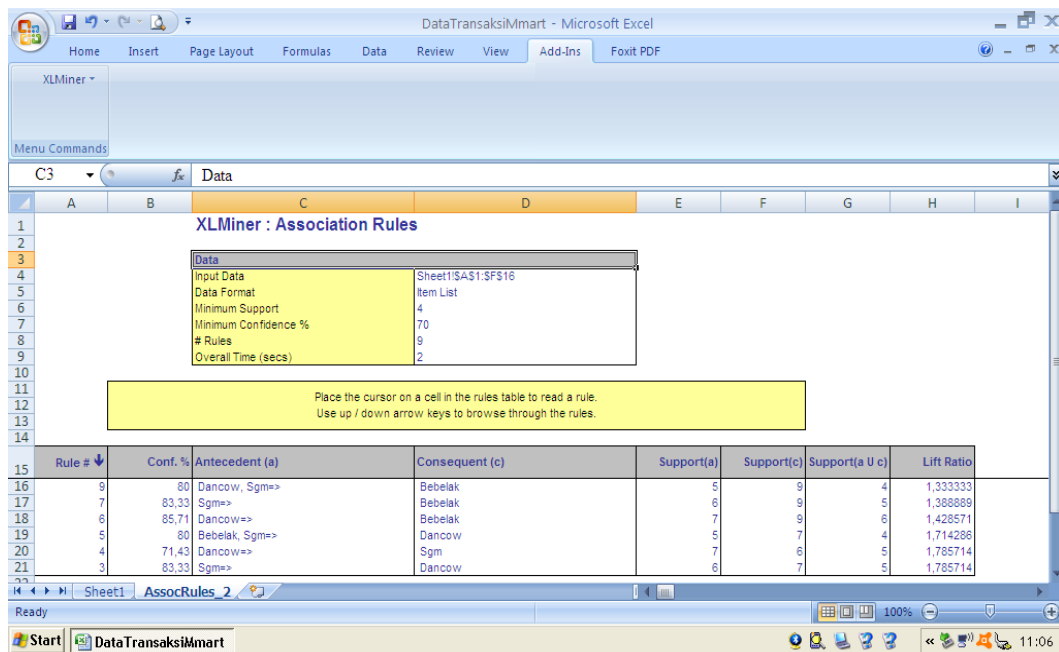


batasan minimum (*Support*) = 30% dan min (*Confidence*) = 70 % adalah {SGM, Bebelac}, {SGM, Dancow}, {Bebelac, Laktogen}, {Bebelac, Dancow}.

**Tabel 6** Hasil aturan asosiasi yang memenuhi *Support* dan *Confidence*

SGM, Bebelac	5/14 =35,7 %	5/6 = 83,3 %
		5/10 = 50 %
SGM, Dancow	5/14 =35,7 %	5/6 = 83,3 %
		5/7 = 71,4 %
Bebelac, Lactogen	5/14 =35,7 %	5/8 = 62,5 %
Bebelac, Dancow	6/14 =42,8%	6/10 = 60 %
		6/7 = 85,7 %

Selanjutnya dilakukan proses pengujian data menggunakan *tools XLMiner3\_2\_4*.



Gambar 2. Hasil pengujian data menggunakan *tools XLMiner*

## Simpulan

Adapun kesimpulan dari hasil penelitian adalah Algoritma *apriori* dapat diterapkan untuk pengambilan keputusan terhadap apa yang berhubungan dengan persediaan barang adalah : Berdasarkan pembahasan yang telah dilakukan susu formula bayi yang banyak diminati atau dibeli oleh konsumen adalah susu SGM, Bebelac, Dancow, Lactogen. Sehingga pemilik mini market bisa memperbanyak stok pada susu tersebut. Selanjutnya penulis juga menyarankan agar dapat membandingkan metode

pengambilan keputusan dengan menggunakan algoritma *apriori* ini dengan teknik lainnya.

### Daftar Pustaka

- Hermawati Fajar A. (2013). *Data Mining*. Andi Offset.
- Kusrini dan Lutfi, Taufik. (2009). *Algoritma Data Mining*, Jakarta, Andi.
- Nugroho Wandu, Rully A. Hendrawan dan Akhmad Mukhlason, (2012). *Pengembangan Sistem Rekomendasi Penelusuran Buku dengan Penggalan Association Rule Menggunakan Algoritma Apriori (Studi Kasus Badan Perpustakaan dan Kearsipan Provinsi Jawa Timur)*. ISSN 2301-9271 Vol. 1.
- Sari Novita Eka, (2013). *Analisa Algoritma Apriori untuk Menentukan Merek Pakaian yang Paling diminati pada Mode Fashion Group Medan*. ISSN 2301-9425 Vol. 4 No. 3.